

**Case Study 1**

**Twin Solar**

**Project Plan Blueprint**

**Rakesh N J**

**Vaishnavi Kanchan**

**Keerti Suryawanshi**

**Edwin Samuel**

**Hanif, Abdullah**

Prof. Dr. Swati Chandna

Applied Data Science and Analytics

**Document Approvals**

|  |  |  |  |
| --- | --- | --- | --- |
| **Role** | **Name** | **Signature©** | **Date** |
| External Stakeholders | Saiful Islam |  |  |
| Internal Stakeholders | Prof. Dr. Swati Chandna |  |  |
| Technical Advisors | Sina Mehraeen,  Kamellia Reshadi |  |  |
| Product Owner |  |  |  |
| Scrum Master |  |  |  |
| Scrum Team |  |  |  |

**Scrum Project Plan**

**Sprint Plan**

**Sprint 0 (April 11th to April 26th): Project Initialization**

**User Story 1:** As a team member we pitched for three topics as per our potential interest and were assigned with one of the case study topics.

**Sprint 1 (April 27th to 12th May): Domain and Data Study**

**User Story 1:** As a Data Scientist I want to deepen my understanding of the microgrid system domain through further research and literature review.

**User Story 2:** As a Data Scientist I want to thoroughly study the provided datasets to identify relevant features and understand their characteristics.

**Sprint 2 (13th May to 28th May): Data Analysis and Database Creation**

**User Story 1:** As a Data Scientist, I want to analyze and investigate the datasets with visualization to identify significant and non-significant parameters.

**User Story 2:** As a Data Engineer I need to create a database and tables for the dataset as per the requirement so that it can be visualized and presented to stakeholder in the form of storytelling.

**Sprint 3 (29th May to 13th June): Feature Selection & Engineering**

**User Story 1:** As a Data Scientist I want to perform feature selection to identify the relevant variables for modeling.

**User Story 2:** As a Data Scientist I want to do feature engineering to improvise the model performance

**Sprint 4 (14th June to 29th June): Baseline Model building**

**User Story 1:** As a Data Scientist I want to build a baseline ML model using the pre-processed data I want to assess the model’s performance to identify the areas of improvement.

**Sprint 5 (30th June to 15th July): Model Execution & Comparison**

**User Story 1:** As a Data Scientist I want to compare and execute different ML models to select the most suitable one for task.

**Sprint 6 (16th July to 30 July): Model Evaluation**

**User Story 1:** As a Data Scientist I want to compare and execute different ML models to select the most suitable one for task.

**User Story 2:** As a Data Scientist I want to identify which parameters contributes more towards the power consumption findings and also need to compare the accuracy with any existing model.

**Sprint 7 (1 August to 15th August): Streaming Visualization**

**User Story 1:** As a Data Engineer I need to visualize the data that has been stored in the database to discover the trends and patterns of power consumption with respect to every quarterly month.

**User Story 2:** As a Data Engineer I need to stream this visualization to provide real time insights.

**Sprint 8 (16th August to 31st August): Documentation**

**User Story 1:**As a Data Scientist I want to illustrate the whole process through project reports and presentations to the stakeholders.

**Sprint 9 (September 1 to September 15th): Backlog**

1. Understanding the business story. Write a document for it.
2. Data Collection, store in Data Lap, take the link from Prof. Swati.
3. Data Management (Edwin)

Data Engineering:

* Cleaning data
* Data collection
* Removing extra columns
* Renaming columns

Data management:

Updating user stories:

Writing the names of each person in the user-stories.